13. TECHNIQUES USED IN BOEING SUBJECTIVE NOISE-RATING STUDIES

By John W. Little, James E. Mabry, Norman A. Yeowart,
The Boeing Company

and S. S. Stevens
Harvard University

INTRODUCTION

For over a decade the effect of jet airplane noise on humans has been evaluated in terms of perceived noise level (PNL). Continued use of the PNL calculation procedure has not prevented further research directed toward obtaining more precise relationships between the physical parameters of the flyover noise and man's response. Corrections to the PNL calculation procedure have resulted from this continuing research for both discrete frequencies (tone above background noise) and duration. However, from the standpoint of flyover-noise-reduction efforts, the most disconcerting finding is that, at times, the amount of noise reduction predicted by the PNL procedure is not matched by human perception. Results from both laboratory and field-test studies of man's response to flyover noise indicate that reductions of 13 to 16 PNdB are required for groups of judges to rate the lower-level noise as one-half the noisiness of the original noise (refs. 1 and 2). These results indicate that, if the aim is to reduce the noise by one-half, the reduction should be on the order of 15 PNdB instead of 10 PNdB, which was the value assumed in the development of the PNL calculation procedure. This inconsistency between subjects' judgments and the assumed rate of change of noisiness played a major role in the selection of the approach used to determine the noise reduction achieved in the National Aeronautics and Space Administration Treated-Nacelle Program.

APPROACH

General

A field-test approach was used, with the main noise stimuli being level flyovers of both an existing 707 airplane and a 707 airplane treated for noise reduction. Annoyance judgments were also obtained for routine airplane flyovers and for random noise presented over loudspeakers.

Six groups of 30 persons judged all the noises at six different listening positions. Three of the listening positions were inside trailers (one directly under the flight path and two at sidelines 1000 and 1500 feet from the flight path). The other three were outside

listening positions adjacent to each of the three trailers. Figure 1 is a schematic drawing of the six positions.

Noises Judged

Judgments were obtained for the annoyance to four groups of stimuli. These four groups are designated as (1) pairs, (2) specials, (3) unscheduled traffic, and (4) loudspeaker noise. Table I shows the number of noise events judged in each of the four groups.

Pairs.- Pairs of level flyovers simulating landing and take-off conditions were used. Table II shows the pairs as planned for the two days of testing. For each condition, the existing airplane was flown repeatedly at a given altitude while the treated airplane was flown at three different altitudes so that the noise levels would be about 5 PNdB less than, equal to, and 5 PNdB greater than the level for the existing airplane. Each of the resulting four sets of three pairs in table II was to be flown twice. On one occasion, the subjects first judged the existing airplane and then the treated airplane, while on the other occasion, the order of presentation for the pair was reversed. One of the main reasons for using pairs of flyovers was to determine the extent of agreement between various engineering calculation procedures and the subjective evaluations, as discussed subsequently in the section "Analysis Techniques." Since it was not possible to complete all of the program for the second day, some of the pairs listed for that day were not flown.

Specials.- Individual flyovers of both the existing and treated airplanes were made at the same altitudes and power settings. These flyovers are referred to as specials.

Unscheduled traffic.- Subjective evaluations were obtained for jet-powered airplanes that routinely use the airport for crew training. This group of stimuli is referred to as unscheduled traffic.

Loudspeaker noise.- At the beginning of each listening session, a 4-second USASI noise at a level of 90 PNdB was presented via loud speakers at all six listening positions. This was the standard to which the subjects were to compare all noises for that session. (USASI noise is random noise with a frequency distribution that peaks at 250 Hz and falls off at -3dB/octave from this peak.) During the session, the 4-second USASI noise was again generated at six different levels during times when flyovers were not occurring. These levels were 70, 75, 80, 85, 90, and 95 PNdB for the inside listening positions and 80, 85, 90, 95, 100, and 105 PNdB for the outside listening positions. The three main reasons for asking the subjects to judge this artificial noise were the following:

(1) Judgment of the artificial noise was a means for determining that the subjects were adequately performing the task.

(2) Since it was not possible to present flyovers on a regular basis, making annoyance judgments of the artificial noise during slack periods kept the subjects at their task.

(3) Subjective evaluations of the same steady-state noise at different levels allowed assessment of the extent that noise-reduction approaches, based on a particular calculation procedure, match judges' ratings of the flyover noise.

Judges and Their Task

The 180 judges for the six listening positions were hired through a state employment agency. There was an equal number of males and females. Applicants were required to respond to a 40-item "Annoyance Questionnaire." Since there were ten items relating to noise in the questionnaire, it was possible to equate the six listening groups on the basis of their reported sensitivity to noise in general. The questionnaire was also used as a kind of screening device. Prior to employment, the questionnaires were examined for ability of the applicants to follow the simple instructions. It was presumed that anyone who could not understand the instructions for the "Annoyance Questionnaire" would also have difficulty in understanding the nature of the noise-judgment task.

The subjects made their judgments under the following instructions, which they were asked to read:

"We are asking you to help us solve a problem concerned with noise: How annoying or disturbing are various kinds of sound when heard in your home? You will be asked to give a score for each sound. First, we will produce a sound whose noisiness score is 10. Use that sound as a standard, and judge each succeeding sound in relation to that standard. For example, if a sound seems twice as noisy as the standard, you will write 20 in the appropriate box on the answer sheet. If it seems only one-quarter as noisy, write 2.5. If it seems three times as noisy, write 30, and so on. Please try to judge each sound carefully, and give it a score that tells how strong the annoyance seems to you. There are no right or wrong answers. The important thing is to say how you rate each of the sounds."

After the subjects had read the instructions, the test director made the following announcement over the loudspeaker systems: "Your rating should be only your own opinion. The test will begin in a moment. Here is the sound which has a noisiness score of 10. Listen to it carefully." To give an idea of the judges' task, the log for the noise judgments of session I is presented as table III.

ENGINEERING CALCULATION PROCEDURES TO BE EVALUATED

Eighteen separate calculation procedures are to be evaluated. The basis for nine of these procedures is the widely used PNL scale. The remaining nine are based on S. S. Stevens' Perceived Loudness Level (Mark VI) from which the PNL approach was originally derived. Each of the two basic engineering calculation procedures is to be evaluated without corrections. In addition, two different tone correction methods utilizing two different methods for identifying the tone will be applied to each basic procedure. Finally, all tone-corrected procedures will be corrected for duration. Thus, the procedures to be evaluated are the following:

(1) Perceived Noise Level (PNL)

(2) Perceived Loudness Level (Mark VI) (PLL)

(3) PNL with tone correction according to FAA Docket No. 9337, Notice No. 69-1

(4) PLL with tone correction according to FAA Docket No. 9337, Notice No. 69-1

(5) PNL with tone correction of Third Revised Draft of Proposed FAA Noise Certification Criteria (May 1, 1967)

(6) PLL with tone correction of Third Revised Draft of Proposed FAA Noise Certification Criteria (May 1, 1967)

(7), (8), (9), and (10) Same as (3), (4), (5), and (6), respectively, but with the tone identified by a four-band averaging technique

(11), (12), (13), (14), (15), (16), (17), and (18) Duration correction to FAA Docket No. 9337, Notice No. 69-1, applied to procedures (3) to (10)

Procedures (11) to (18) yield various forms of Effective Perceived Noise Level. Procedure (11) is recommended by the FAA in Docket No. 9337, Notice No. 69-1.

ANALYSIS TECHNIQUES

The data were collected in such a manner that two analysis approaches could be utilized. The two approaches are the method of constant-stimulus differences (sometimes referred to as the method of pair comparisons) and the method of direct magnitude estimation.

Constant-Stimulus Differences

The aim of this method is to determine the extent that the subjective evaluations agree with a particular calculation procedure. Figure 2 shows some expected results for pairs 4, 5, and 6 from judgments made at the positions directly under the flight path.

The premise and conclusion for this approach is that if two different sounds are judged equally annoying, an acceptable engineering calculation procedure assigns the identical real number to both sounds.

Figure 2 shows a case of good agreement between the subjective judgments and the calculation procedure since the discrepancy is less than 1 PNdB. All eighteen procedures are to be evaluated by using the available sets of pairs, and the procedure for which the agreement is best becomes a likely candidate for describing the amount of noise reduction achieved.

Direct Magnitude Estimation

The magnitude-estimation method, refined by S. S. Stevens of Harvard, has been used widely as a method of relating human response to a physical stimulus. Data from many studies indicate that sensation is a power function of the stimulus. Consequently, in log-log plots a linear equation is obtained for the relationship between the judgments and the physical evaluation of the noise.

Figure 3 shows a plot of one set of expected results. Note that the rate of change of annoyance for the airplane flyovers is less than that for the USASI noise. This prediction is made on the basis of the previously mentioned studies (refs. 1 and 2).

CONCLUDING REMARKS

The findings of this study must be analyzed by both the constant-stimulus-differences technique and the magnitude-estimation technique for each engineering calculation procedure. Subsequently, the results must be compared with previous relevant sets of similar data that describe human response to airplane flyover noise. It is naturally expected that good agreement will exist between the two analysis techniques since they both draw from the same pool of data. However, when subject variability is high or when the sounds being judged are unique enough to bias responses so that they are different than expected, the two methods may produce different findings.

At this time there seem to be two very distinct advantages to the described test method for field testing of subjects. First, there is the strong advantage that at least two different analysis techniques can be used on the same set of data. Second, the method has excellent practical advantages since a precise test format is not essential. Testing with aircraft in flight depends on having good weather and uninterrupted aircraft performance, neither of which can be guaranteed. The described method allows for test-plan deletions or additions at will with no significant loss in data if only a few conditions of a set are missed.

Finally, it must be remembered that most of the past work relating human response to aircraft noise has been done in laboratory studies, which are usually simpler and less expensive than field tests of the type described in this report. Thus, there is a natural concern that correlation be established between field and laboratory findings. All sounds judged in the present field test are stored on magnetic tape, and thus a source of stimulus material is provided for subsequent laboratory investigation.

REFERENCES

1. Broadbent, D. E.; and Robinson, D. W.: Subjective Measurements of the Relative Annoyance of Simulated Sonic Bangs and Aircraft Noise. J. Sound Vib., vol. I, no. 2, Apr. 1964, pp. 162-174.

2. Bishop, Dwight E.: Judgments of the Relative and Absolute Acceptability of Aircraft Noise. J. Acoust. Soc. Amer., vol. 40, no. 1, July 1966, pp. 108-122.

TABLE I.- NOISE EVENTS

Noise group	Number of noise events		
	First day	Second day	Total
Pairs	12	10	22
Specials	8	3	11
Unscheduled traffic	3	11	14
Loudspeaker noises:			
Standards*	1 + 5	1 + 5	2 + 10
Other levels	25	25	50

*The standard was identified for the subject only one time each day but was presented five more times.

TABLE II.- PAIRS OF EXISTING AND TREATED AIRPLANES

[E = Existing airplane; T = Treated airplane]

First day			Second day								
Simulated landings with 5000-lb thrust			Simulated landings with 5000-lb thrust			Simulated take-offs with maximum thrust			Simulated take-offs with cutback thrust		
Airplane	Altitude (ft)	PNL (PNdB)	Airplane	Altitude (ft)	PNL (PNdB)	Airplane	Altitude (ft)	PNL (PNdB)	Airplane	Altitude (ft)	PNL (PNdB)
E	2000	99	E	2000	99	E	2100	108	E	3400	98
T	1150	94	T	1150	94	T	2100	103	T	2100	93
T	850	99	T	700	99	T	1600	108	T	1600	98
T	650	104	T	450	104	T	1100	113	T	1100	103

TABLE III.- LOG FOR NOISE JUDGMENTS OF SESSION I

[E = Existing airplane; T = Treated airplane]

No.	Code (*)	Notes	Time	No.	Code (*)	Notes	Time
0	3u	Standard at 90 PNdB	2:40	27	5	E (altitude of 2040 ft)	3:20
1	3u	90 PNdB		28	5	T (altitude of 875 ft)	3:23
2	4u	95 PNdB		29	6u	105 PNdB	
3	1u	80 PNdB		30	2u	85 PNdB	
4	1u	80 PNdB		31	4u	95 PNdB	
5	(8)	E (special landing)	2:47	32	4u	95 PNdB	
6	1u	80 PNdB		33	1	T (altitude of 1255 ft)	3:36
7	DC-8	No recording	2:48	34	1	E (altitude of 1980 ft)	3:38
8	2u	85 PNdB		35	3u	90 PNdB	
9	6u	105 PNdB		36	1u	80 PNdB	
10	2	T (altitude of 870 ft)	2:49	37	1u	80 PNdB	
11	2	E (altitude of 2100 ft)	2:56	38	5u	100 PNdB	
12	4u	95 PNdB		39	6u	105 PNdB	
13	3u	90 PNdB		40	4	E (altitude of 2100 ft)	3:44
14	6u	105 PNdB		41	4	T (altitude of 1170 ft)	3:47
15	(7)	T (special landing)	2:59	42	3u	90 PNdB	
16	DC-8	Landing		43	5u	100 PNdB	
17	5u	100 PNdB		44	5u	100 PNdB	
18	2u	85 PNdB		45	3u	90 PNdB	
19	2u	85 PNdB		46	6	E (altitude of 2180 ft)	3:54
20	3	T (altitude of 640 ft)	3:11	47	6	T (altitude of 710 ft)	3:56
21	3	E (altitude of 2000 ft)	3:14	48	(9)	E (1000-ft alt.; max. take-off thrust)	4:02
22	DC-8	Landing	3:15	49	(10)	T (1000-ft alt.; max. take-off thrust)	4:03
23	6u	105 PNdB		50	(11)	E (1000-ft alt.; cutback thrust)	4:11
24	5u	100 PNdB		51	(12)	T (1000-ft alt.; cutback thrust)	4:13
25	4u	95 PNdB		52	(13)	E (700-ft alt.; actual landing)	4:19
26	2u	85 PNdB		53	(14)	T (700-ft alt.; actual landing)	4:20

*u refers to USASI noise; all USASI noise durations are 4 seconds.

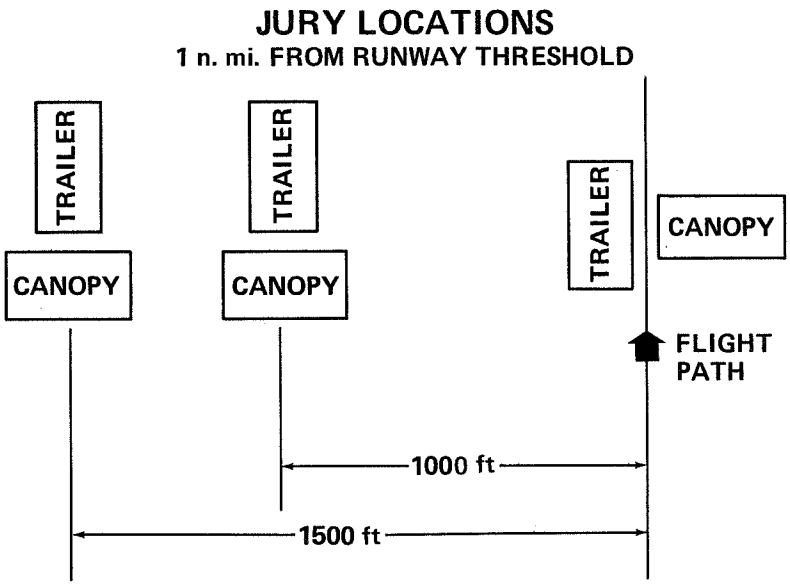

Figure 1

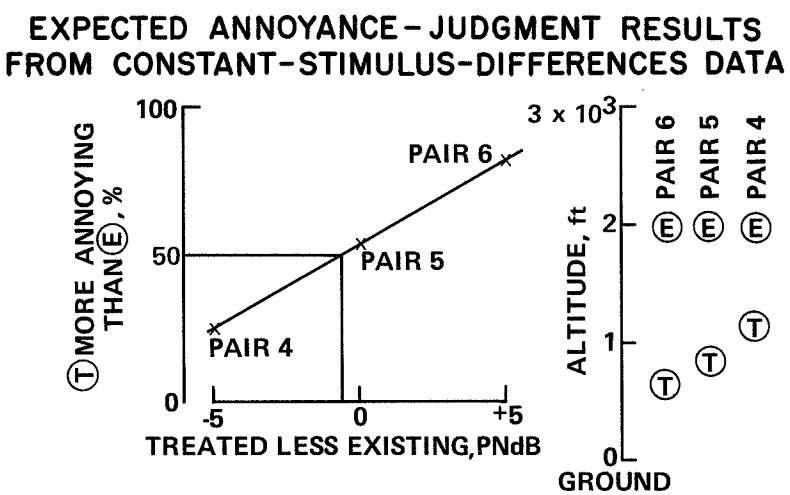

Figure 2

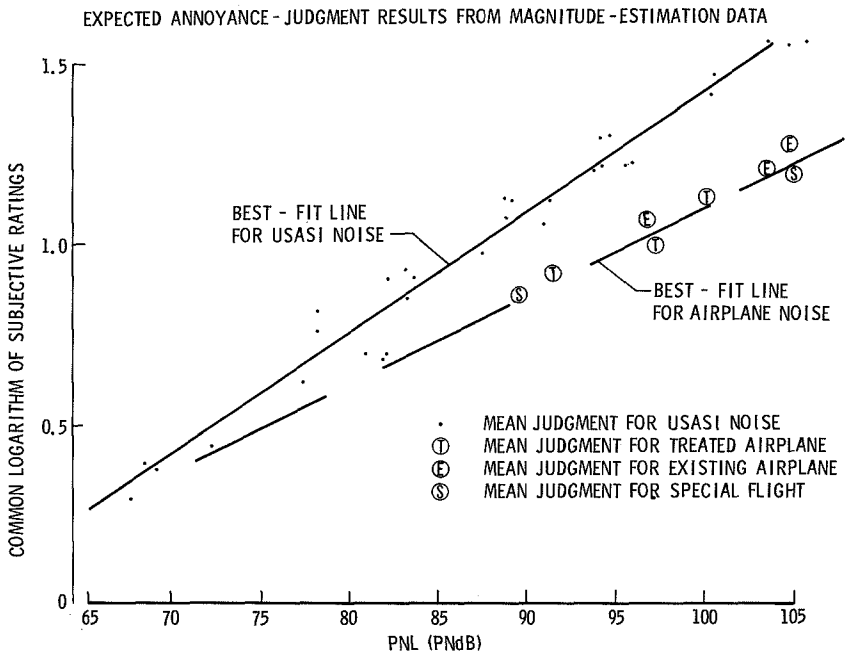
EXPECTED ANNOYANCE - JUDGMENT RESULTS FROM MAGNITUDE - ESTIMATION DATA
COMMON LOGARITHM OF SUBJECTIVE RATINGS
1.5
1.0
0.5
0
65
70
75
80
85
90
95
100
105
PNL (PNdB)
BEST - FIT LINE FOR USASI NOISE
BEST - FIT LINE FOR AIRPLANE NOISE
MEAN JUDGMENT FOR USASI NOISE
MEAN JUDGMENT FOR TREATED AIRPLANE
MEAN JUDGMENT FOR EXISTING AIRPLANE
MEAN JUDGMENT FOR SPECIAL FLIGHT

Figure 3